Exploring Fundamental Frequency Characteristics of Australian Adolescents with and without Depression in the Future Proofing Study Corpus

Tünde Szalay[1], Brian Stasak[1,2], Kate Maston[2], Aliza Werner-Seidler[2], Mark Larsen[2,3]

[1]School of Elect. Eng. & Telecomm., [2]Black Dog Institute, [3]Centre for Big Data Research in Health
University of New South Wales

t.szalay@unsw.edu.au

Abstract

Lower mean and reduced fundamental frequency ($F0$) variation are speech-related markers of depression in adults. Using the Future Proofing Corpus, $F0$ differences in contours, variation, and creakiness were examined in read aloud sentences produced by 14 Australian adolescent female speakers with and without depression. Contrary to known $F0$ markers of depression, the current study found no evidence for speakers with depression producing lower or less varied $F0$ than speakers without depression, potentially due to other factors (e.g., age, puberty) masking depression-related $F0$ differences. However, speakers with depression were significantly more likely to produce creaky voice than their peers without depression.

Index Terms: Australian English, corpus phonetics, creaky voice, $F0$, mental health, paralinguistics

1. Introduction

Depression has a prevalence rate of 15% among Australian adolescents [1]. The Future Proofing Study (FPS) examines factors associated with mental health conditions, including depression, during adolescence using big data [1]. The FPS collected demographic and mental health information using surveys, as well as audio recordings to capture speech markers of mental health conditions [1, 2]. The FPS corpus contains speech data suitable for acoustic analysis of changes in speech associated with depression.

Changes in speech, such as lower speech rate, increased hesitation markers, and lower and less varied intonation, are frequent speech markers of depression [3, 4]. In adults, clinical literature describes the voice of speakers with depression as “low”, “monotonous”, and “toneless” due to listeners auditory-impressionistic perception of patients with depression as having lower and less varied pitch [4]. In line with the perceived low pitch of speakers with depression, overall fundamental frequency ($F0$) was quantitatively found to reduce with depression severity in adults [5, 6]. Similarly to adults, adolescent male speakers with depression produced lower mean $F0$ compared to adolescents without depression when discussing a series of topics, such as event planning or conflict resolution [7]. In contrast, adolescent female speakers with depression produced higher $F0$ values compared to their peers without depression, but only when discussing event planning [7].

Supporting perceptual clinical observations on depressed voice being “monotonous”, speakers with depression show relatively narrower range, less variation, and slower rate of changes in $F0$ in acoustic studies [5, 8]. Speakers with depression produce less $F0$ variation based on context and topic compared speakers without depression [5, 8]. For example, speakers with depression have less $F0$ variability differentiating happy, sad, or angry speech [5]. Speakers with more severe depressive symptoms produce less $F0$ variation to distinguish clear speech from conversational speech compared to speakers with less severe symptoms [8]. Reduced $F0$ variation can be explained by increased muscle tension in the larynx caused by psychomotor retardation, i.e., the slowing of thought and reduction of physical movements accompanying depression [4]. However, increasing $F0$ range and velocity with increasing depression severity has also been found by acoustic studies [9].

Inconsistent results are attributed to $F0$ being affected by multiple factors besides depression, such as feeling agitated or anxious [7, 10]. For example, patients produced lower $F0$ after depression treatment compared to pre-treatment due to feeling more relaxed post-treatment [10]. Despite the inconsistent acoustic results, perceived pitch may be considered during clinical assessments, and $F0$ measurements have contributed to the development of speech-based automatic depression screening algorithms achieving approximately 70% accuracy [3, 11].

The majority of studies showing $F0$ differences associated with depression were carried out on American English adult [4] and adolescent speech [7, 11]. Despite potential accent differences in $F0$, automatic depression detection algorithms can successfully classify speech as depressed or non-depressed using $F0$ characteristics of Australian English (AusE) speech [12, 13].

The FPS corpus provides self-reported symptoms of depression combined with mobile phone speech data suitable for developing automatic depression detection algorithms for a population that may be difficult to access: young adolescents reporting symptoms of depression. To improve validity and interpretability of such algorithms, we examined $F0$ patterns of Australian adolescents with and without depression in the FPS corpus. We tested whether $F0$ markers of depression present in adults are also present in the young adolescents in the FPS corpus, hypothesising that AusE-speaking adolescents reporting symptoms of depression would produce lower $F0$ and reduced $F0$ variation.

2. Methods

2.1. Corpus description

The FPS corpus contains survey data from 6,388 participants and 14,498 sound files elicited from 1,062 year 8 students recruited from 134 secondary schools across Australia and collected across multiple timepoints [2]. Participants undertook the data collection activities on their own time using their smartphones with no obligation to complete the speech recordings, resulting in the different number of participants submitting survey- and audio data.

Survey data included demographic information, such as participants' age, linguistic- and cultural background; health information concerning disabilities, drug-consumption and COVID-19; and a series of mental health symptoms. Symptoms of depression were measured using the Patient Health Questionnaire for Adolescents (PHQ-A) [14]. PHQ-A ranges from 0 to 27, with higher scores indicating more severe depression; PHQ-A $\geq$ 15 was used as the threshold for caseness [1].

Two protocols were used for speech elicitation. The first protocol included the diadochokinetic /pataka/ task, the Harvard sentences [15], the Rainbow passage [16] and was used from August 2019 to April 2021. The second protocol included a sustained /aː/; diadochokinetic /pataka/; a prosodic sentence set designed to elicit different intonation contours associated with describing movement; an affective sentence set with positive, negative, and neutral valence designed to elicit emotions; open-ended question; categorisation; and a creative task asking participants to list homonyms. The second protocol was used from April 2021 onwards, until the end of data collection.

Students submitted survey and speech data in a self-timed, unsupervised manner at a baseline timepoint and at eight follow-up points [2]. Speech and survey data were submitted to two different servers and linked via timestamps. The data collection method allowed for delayed submission of speech samples relative to the survey data. That is, depression severity may have been different at the time of audio recording than what was reported in the survey.

2.2. Corpus data extraction

To reduce changes in $F0$ associated with participants going through puberty during real-time data collection, only sound files submitted within 31 days of baseline survey data collection were considered in the analysis. To reduce $F0$ variation between speakers due to non-depression-related factors, only Australian-born cisgender speakers who reported speaking English at home and reported no disabilities were included.

To select the most suitable speech tasks, we considered constraints of phonetic and clinical validity. Only read speech tasks were considered as they had a uniform format, length, and contents. For clinical validity, we only considered sentence- or passage reading tasks. To increase the likelihood of capturing $F0$ differences, we selected the tasks designed to elicit large $F0$ range and variation. The prosodic sentence set was designed to elicit $F0$ variation via the juxtaposition of contrasting motion- and speed-related words (e.g., *The hare quickly went down the hill* vs. *The bear slowly went up the hill*), consisting of 24 paired sentences. Out of the 12 pairs, one pair (two sentences) were assigned randomly to each participant. The affective sentence set contained 10 positive, 10 negative, and 10 neutral sentences, designed to elicit a variety of $F0$ patterns due to $F0$ being a key marker of emotions [17]. Out of the 30 sentences, one positive, one negative, and one neutral sentences were assigned randomly to each participant. The prosodic and affective sentence tasks combined provide five sentences per participant.

Out of the 1,061 speakers submitting audio files, 781 submitted audio at baseline, 574 met speaker inclusion criteria, and 69 submitted the prosodic and affective sentence tasks. Of the 69 speakers, 15 (M = 5, F = 10) reached the threshold for caseness of depression. Male speakers were excluded due to the low number of male speakers with depression in the sample. Three female speakers with depression were excluded due to low audio quality, and seven were included. The seven included speakers with depression (F = 7) had a mean age of 13.46 years and reported diverse sexualities (lesbian/bi-/pansexual = 5, heterosexual = 2). As lower $F0$ is expected to mark both age and depression [18], the 7 adolescents with depression were matched with 7 female adolescents without depression according to age (mean = 13.34 years). As lesbian sexual orientation may be indexed in other languages by lowered and less varied $F0$ [19, 20], the markers of depression under investigation, the 7 adolescents without depression matched the group with depression according to sexuality (diverse = 5, heterosexual = 2).

In line with analysis of the survey responses of the baseline cohort (6,388 participants), speakers with depression were more likely to be female than male and more likely to identify with diverse sexualities than identify as heterosexual among participants submitting speech data [1].

2.3. $F0$ analysis

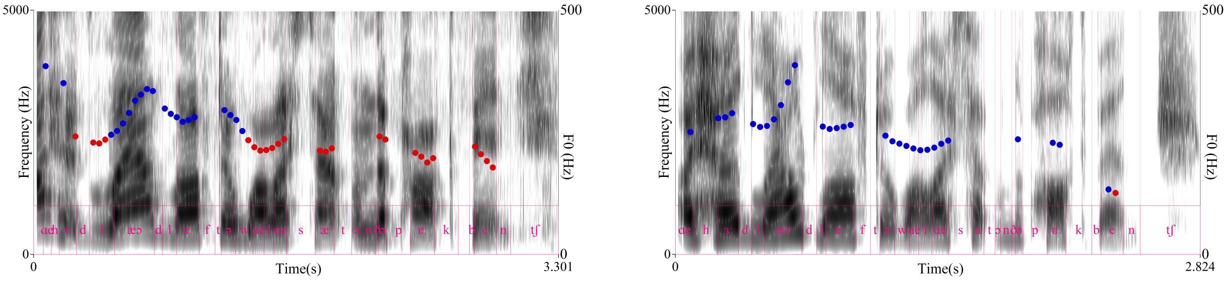
A total of 14 (speakers) $\times$ 5 (sentences) = 70 sentences were included in the analysis. $F0$ analysis followed the protocol recommended for semi-automatic $F0$ estimation [21]. Recorded sentences were force-aligned automatically using the Montreal Forced Aligner with a British pronunciation dictionary and a multi-accented acoustic model; manual correction was not completed to improve replicability [22, 23]. $F0$ values were estimated automatically and overlaid on spectrograms for visual inspection at every 37.5 ms between 20 Hz to account for creaky voice and 600 Hz to account for adolescent female speakers having high $F0$ [21, 24]. $F0$ estimates were time-aligned with the automatic phonetic segmentation; $F0$ values estimated during non-sonorant segments were excluded [21]. $F0$ variation was captured as the absolute value of $F0$ difference between $F0$ measured at the end of a sonorant interval and $F0$ measured at the start of the following sonorant interval.

2.3.1. *Automatic creak detection*

As young AusE speakers are likely to produce creaky voice, automatic creak detection was applied to the data by visualising each speaker's $F0$ distribution using histograms [21, 25, 26]. Creaky voice was identified in bimodal distributions,by the presence of a second peak with a lower mode, separated from the higher peak of modal voice by an anti-mode [26, 21].

The number of modes was determined using the excess mass test in the *antimode* library of R and visual inspection [27, 28, 29]. When both excess mass test and visual inspection indicated bimodal $F0$ distribution, the speaker was considered to have produced creaky and modal voice. When the excess mass test indicated bimodal $F0$ distribution, but visual inspection suggested that the $F0$ distribution was right-skewed rather than bimodal, the speaker was considered to have produced modal voice. When neither the excess mass test nor the visual inspection indicated bimodal distribution, the speaker was considered to have produced modal voice. $F0$ values produced below the anti-mode in bimodal distributions were rated as creaky (Figure 1a). For all other speakers, $F0$ values below 132 Hz threshold were rated as creaky (Figure 1b) [26].

Average $F0$ of modal voice in the current FPS data produced by 12-14-year-old female speakers fell between that of 11-year-old prepubescent female speakers and 18-29-year-old adult female speakers (11-year-old = 248 Hz; 12-14-year-old = 237 Hz, 18-29-year-old = 206 Hz) [18, 26].



(a) *Predominantly creaky voice produced by a speaker with depression.* 60% of the F0 estimates identified as creaky and 40% as modal.

(b) *Predominantly modal voice produced by a speaker without depression.* 2.9% of F0 estimates identified creaky and 97.1% as modal.

Figure 1: *The sentence* I heard loud laughter while I sat on the park bench *produced by two speakers, overlaid with automatic segmentation (pink) and* F0 *contour (creaky: red, modal: blue).*

2.4. Statistical analysis

As automatic creak detection indicated a potential difference in creak between speakers with and without depression, $F0$ produced during modal and creaky voice were included in the statistical analysis. In addition, differences in the proportion of creaky voice between speakers with and without depression were examined statistically, despite not being hypothesised.

To test $F0$ differences between speakers with and without depression, one linear mixed model and two generalised linear mixed models were constructed. The response variable $F0$ was modelled using a linear mixed model with Gaussian family. The response variable $F0$ Variation was modelled using a generalised linear mixed model with Gamma family, as $F0$ Variation had a right-skewed distribution. The presence of creak was modelled using generalised mixed model with binomial family with the response variable Creaky (modal $F0$ coded as 0, creaky $F0$ coded as 1). In all models, the independent variable was Depressed (categorical, comparing "With Depression" to the baseline "No Depression") with a by-participant random effect to account for intraspeaker variation, such as age, puberty, and sexual orientation. All models were constructed using the *lme4* library in R; p-values were calculated with the *lmerTest* library using Satterthwaite's degrees of freedom method [30, 31].

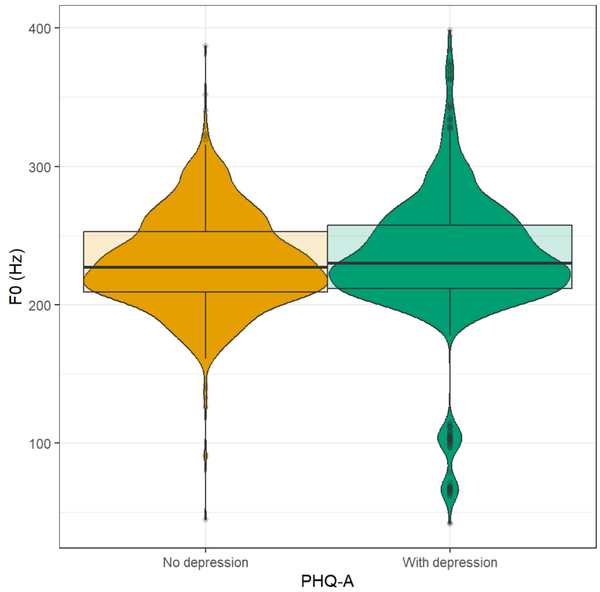


Figure 2: F0 *produced by speakers with and without depression; random effect of speaker not visualised.*

3. Results

Speakers with and without depression did not show a significant difference in $F0$ and $F0$ variation ($F0$: $\beta = -1.45$ Hz, $t_{12.71} = -0.11, p = 0.911$, Fig. 2; $F0$ variation: $\beta = -0.01$ Gamma-transformed Hz, $t_{0.01} = -1.31, p = 0.19$, Fig. 3).

Speakers with depression produced a significantly higher proportion of creaky voice compared to speakers without depression ($\beta = 3.92\%, z_{1.59} = 2.46, p = 0.0138$, Fig. 4). Four out of seven speakers with depression produced bimodal $F0$ distribution and three produced unimodal distribution. Out of the three speakers with depression producing unimodal distribution, one produced $F0$ values below the 132 Hz threshold rated as creaky, and two did not produce $F0$ values rated as creaky at all. All seven speakers without depression produced unimodal $F0$ distribution, with two producing values below the threshold of creaky voice, and five producing $F0$ rated as modal.

4. Discussion

The study aimed to test if $F0$ markers of depression were present in the FPS corpus of Australian adolescents. Adolescents reporting symptoms of depression were hypothesised to produce lower and less varied $F0$. Contrary to the hypotheses, we did not find differences in overall $F0$ or in $F0$ variation between speakers with and without depression. That is, lowered and less varied $F0$ symptoms were either not present in the current sample of the FPS corpus or were too small to discover.

Changes in $F0$ accompanying depression are expected have small effect size, as shown by the inconsistent $F0$ differences in the literature [4]. In our study, small effects of depression could have been masked by other factors, such as more robust puberty related differences. Although groups with and without depression were balanced for age, adolescents may not go through puberty at the same time, thus, one group may have more speakers who have reached puberty than the other. The potential effects of sexual orientation may also mask $F0$ symptoms of depression if adolescents reporting diverse sexual orientation use lower and less varied $F0$ [19, 20]. However, effects of sexual orientation cannot be reliably tested in the current sample.

Exploratory analysis revealed that adolescent female speak-

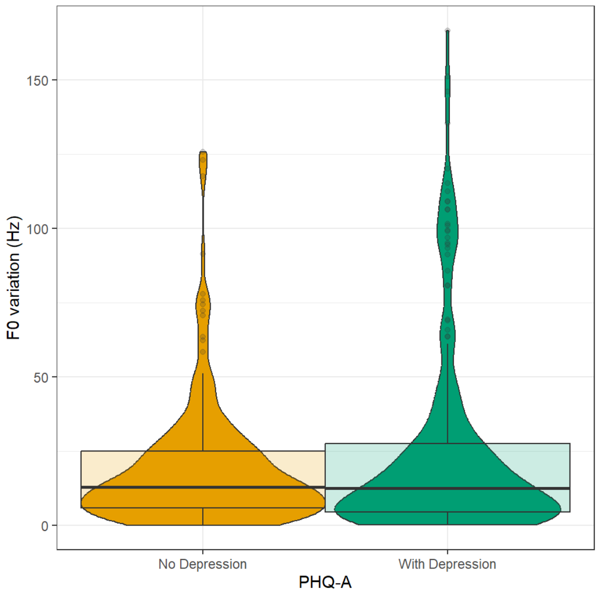


Figure 3: *F0 variation produced by speakers with and without depression; random effect of speaker not visualised.*

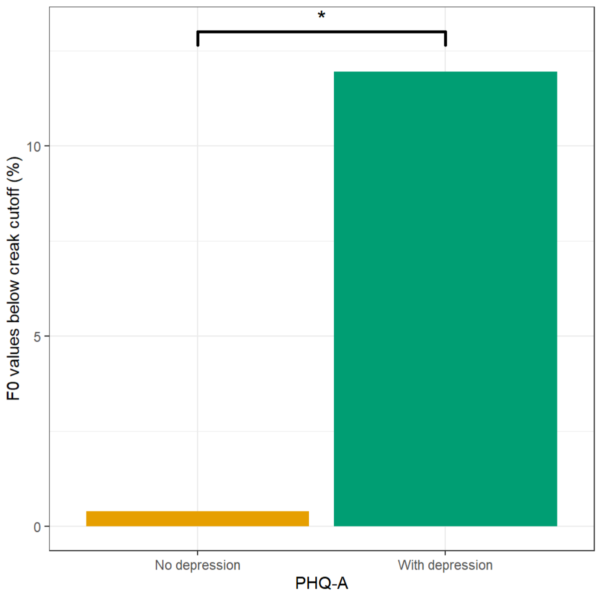


Figure 4: *Percentage of* F0 *estimates below the cutoff-point of creak relative to the total number of* F0 *estimates per group; random effect of speaker was not visualised.*

ers with depression produced a higher proportion of creaky voice compared to the speakers not reaching the threshold of depression. That is, depression may be marked by intra-speaker variation as speakers lower their $F0$ relative to themselves using creak rather than relative to their peers without depression. Low open quotient, another characteristic of creaky voice, has also been identified as a speech marker of depression [32, 33]. In contrast, when voice quality was captured using harmonics-to-noise ratio and jitter, increased depression severity was marked by aspiration and breathiness, indicating breathy, rather than creaky voice [9]. Results on voice quality as a marker of depression are difficult to compare between studies due to lack of standardisation in extraction techniques [4]. Future research may examine depression and creak using a variety of acoustic characteristics of creaky voice [33].

As $F0$ produced during creaky voice is lower than $F0$ produced during modal voice [33], increased creak among speakers with depression could have resulted in the expected overall lower $F0$. Despite the increase in creak, $F0$ lowering was not observed, potentially due to the relatively low proportion of creaky to modal voice (Figure 4). Low prevalence of creak is attributed to speaker- and task effects. The current sample only contained female speakers who are less likely to produce creaky voice than male speakers [34]. Creak is often used a socio-indexical marker [34]; as speech was elicited via read sentences, the speakers might not have produced such markers.

The lack of difference in overall $F0$ and $F0$ variation is inconsistent with $F0$ differences contributing to the successful classification of speech as belonging to speakers with and without depression [12, 13]. However, increased presence of creaky voice may also contribute to automatic depression detection via the acoustic correlates of creak, such as low and irregular $F0$, high noise, or high harmonics-to-noise ratio [33]. Future research developing automatic depression screening algorithms using the FPS corpus is required to identify the most suitable speech features for automatic depression classification in the FPS corpus. As automatic classification algorithms consider multiple features rather than just $F0$, $F0$ and creak may be suitable to classify speakers as depressed or not depressed when combined with other speech factors.

The main limitation of the current study is the small sample size. While the overall corpus is large, containing 14,498 sound files from 1,062 participants, only 70 sentences from 14 participants were analysed in the current study. The number of suitable speakers reduced substantially due to extracting only the affective sentences and prosodic sentences tasks submitted at baseline. As baseline data were collected from August 2019 to March 2022, and the second protocol containing affective and prosodic sentences was introduced in April 2021, fewer participants submitted the affective- and prosodic sentences tasks compared to speech tasks from the first protocol, such as The Rainbow passage. Although design and selection of the affective and prosodic sentences tasks were motivated by capturing $F0$ variation, a key speech symptom of depression, the low number of speakers submitting these tasks is likely to have contributed to the null results. Future research may explore the use of other speech tasks, for instance The Rainbow passage, to examine the links between $F0$, $F0$ variation, creak, and depression in the FPS corpus.

5. Conclusion

We analysed $F0$ differences between AusE-speaking adolescent female speakers with and without symptoms of depression using the Future Proofing Study corpus. The speakers included in this study did not exhibit differences in overall $F0$ or $F0$ variation. It is possible that the relatively small effects of depression on $F0$ were masked by other, more robust factors, such as puberty. Speakers with depression; however, produced a higher proportion of creaky voice. Future work will expand the current analysis to test differences in the use of creaky and modal voice between speakers with and without depression using more data from the Future Proofing Study Corpus.

6. Acknowledgements

The Future Proofing Study was approved by the University of New South Wales (NSW) Human Research Ethics Committee (HC180836), the State Education Research Applications Process for the NSW Department of Education (SERAP2019201), and relevant Catholic Schools Dioceses across Australia. This project was supported by the NHMRC GNT1197074 and GNT1152952 grants to Aliza Werner-Seidler and NHMRC GNT1138405 and GNT1155614 grants to Helen Christensen. The authors thank Helen Christensen for her work, and Hannah White and Timothy Shea for their comments.

7. References

[1] A. Werner-Seidler, K. Maston, A. L. Calear, P. J. Batterham, M. E. Larsen, M. Torok, B. O'Dea, K. Huckvale, J. R. Beames, L. Brown *et al.*, "The future proofing study: Design, methods and baseline characteristics of a prospective cohort study of the mental health of Australian adolescents," *International Journal of Methods in Psychiatric Research*, vol. 32, no. 3, p. e1954, 2022.

[2] A. Werner-Seidler, K. Huckvale, M. E. Larsen, A. L. Calear, K. Maston, L. Johnston, M. Torok, B. O'Dea, P. J. Batterham, S. Schweizer *et al.*, "A trial protocol for the effectiveness of digital interventions for preventing depression in adolescents: The future proofing study," *Trials*, vol. 21, no. 1, pp. 1–21, 2020.

[3] K. R. Scherer, "Vocal assessment of affective disorders," in *Depression and expressive behavior*, J. D. Maser, Ed. New York: Routledge, 1987, pp. 57–82.

[4] N. Cummins, S. Scherer, J. Krajewski, S. Schnieder, J. Epps, and T. F. Quatieri, "A review of depression and suicide risk assessment using speech analysis," *Speech communication*, vol. 71, pp. 10–49, 2015.

[5] Z. Breznitz, "Verbal indicators of depression," *The Journal of general psychology*, vol. 119, no. 4, pp. 351–363, 1992.

[6] J. Wang, L. Zhang, T. Liu, W. Pan, B. Hu, and T. Zhu, "Acoustic differences between healthy and depressed people: a cross-situation study," *BMC psychiatry*, vol. 19, pp. 1–12, 2019.

[7] A. Y. Hussenbocus, M. Lech, and N. B. Allen, "Statistical differences in speech acoustics of major depressed and non-depressed adolescents," in *2015 9th International Conference on Signal Processing and Communication Systems (ICSPCS)*. IEEE, 2015, pp. 1–7.

[8] H. Yi, R. Smiljanic, and B. Chandrasekaran, "The effect of talker and listener depressive symptoms on speech intelligibility," *Journal of Speech, Language, and Hearing Research*, vol. 62, no. 12, pp. 4269–4281, 2019.

[9] T. F. Quatieri and N. Malyska, "Vocal-source biomarkers for depression: A link to psychomotor activity." in *Interspeech*, vol. 2, 2012, pp. 1059–1062.

[10] F. Tolkmitt, H. Helfrich, R. Standke, and K. R. Scherer, "Vocal indicators of psychiatric treatment effects in depressives and schizophrenics," *Journal of communication disorders*, vol. 15, no. 3, pp. 209–222, 1982.

[11] L.-S. A. Low, N. C. Maddage, M. Lech, L. B. Sheeber, and N. B. Allen, "Detection of clinical depression in adolescents' speech during family interactions," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 3, pp. 574–586, 2010.

[12] S. Alghowinem, R. Goecke, M. Wagner, J. Epps, G. Parker, M. Breakspear *et al.*, "Characterising depressed speech for classification," 2013.

[13] B. Stasak, J. Epps, and R. Goecke, "Automatic depression classification based on affective read sentences: Opportunities for text-dependent analysis," *Speech Communication*, vol. 115, pp. 1–14, 2019.

[14] J. G. Johnson, E. S. Harris, R. L. Spitzer, and J. B. Williams, "The patient health questionnaire for adolescents: validation of an instrument for the assessment of mental disorders among adolescent primary care patients," *Journal of Adolescent Health*, vol. 30, no. 3, pp. 196–204, 2002.

[15] E. H. Rothauser, "IEEE recommended practice for speech quality measurements," *IEEE Transactions on Audio and Electroacoustics*, vol. 17, no. 3, pp. 225–246, 1969.

[16] G. Fairbanks, "Voice and articulation drillbook," *(No Title)*, 1960.

[17] T. Johnstone and K. R. Scherer, "Vocal communication of emotion," *Handbook of emotions*, vol. 2, pp. 220–235, 2000.

[18] P. A. Busby and G. L. Plant, "Formant frequency values of vowels produced by preadolescent boys and girls," *The Journal of the Acoustical Society of America*, vol. 97, no. 1, pp. 2603–2606, 1995.

[19] B. Moonwomon-Baird, "Toward a study of lesbian speech," *Queerly phrased: Language, gender, and sexuality*, pp. 202–213, 1997.

[20] J. Van Borsel, J. Vandaele, and P. Corthals, "Pitch and pitch variation in lesbian women," *Journal of Voice*, vol. 27, no. 5, pp. 656–e13, 2013.

[21] H. White, J. Penney, A. Gibson, A. Szakay, and F. Cox, "Evaluating automatic creaky voice detection methods," *The Journal of the Acoustical Society of America*, vol. 152, no. 3, pp. 1476–1486, 2022.

[22] M. McAuliffe and M. Sonderegger, "English (UK) MFA dictionary v3.0.0," https://mfa-models.readthedocs.io/pronunciationdictionary/English/English(UK)MFAdictionaryv3_0_0.html, Tech. Rep., Feb 2024.

[23] ——, "English MFA acoustic model v3.0.0," https://mfa-models.readthedocs.io/acoustic/English/EnglishMFAacousticmodelv3_0_0.html, Tech. Rep., Feb 2024.

[24] P. Boersma and D. Weenink, "Praat 6.4.01,," 2023. [Online]. Available: http://www.fon.hum.uva.nl/praat/

[25] J. Penney, F. Cox, and A. Szakay, "Glottalisation, coda voicing, and phrase position in Australian English," *The Journal of the Acoustical Society of America*, vol. 148, no. 5, pp. 3232–3245, 2020.

[26] Y. Leung, J. Oates, V. Papp, and S.-P. Chan, "Speaking fundamental frequencies of adult speakers of Australian English and effects of sex, age, and geographical location," *Journal of Voice*, vol. 36, no. 3, pp. 434–e1, 2022.

[27] J. Ameijeiras-Alonso, R. M. Crujeiras, and A. Rodriguez-Casal, "Multimode: An R package for mode assessment," *arXiv preprint arXiv:1803.00472*, 2018.

[28] J. Ameijeiras-Alonso, R. M. Crujeiras, and A. Rodríguez-Casal, "Mode testing, critical bandwidth and excess mass," *Test*, vol. 28, pp. 900–919, 2019.

[29] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2023. [Online]. Available: https://www.R-project.org/

[30] D. Bates, M. Mächler, B. Bolker, and S. Walker, "Fitting linear mixed-effects models using lme4," *Journal of Statistical Software*, vol. 67, no. 1, p. 1–48, 2015.

[31] A. Kuznetsova, P. B. Brockhoff, and R. H. B. Christensen, "lmerTest package: Tests in linear mixed effects models," *Journal of Statistical Software*, vol. 82, no. 13, p. 1–26, 2017.

[32] S. Scherer, G. Stratou, J. Gratch, and L.-P. Morency, "Investigating voice quality as a speaker-independent indicator of depression and ptsd." in *Interspeech*, 2013, pp. 847–851.

[33] P. A. Keating, M. Garellek, and J. Kreiman, "Acoustic properties of different kinds of creaky voice." in *ICPhS*, vol. 1, 2015, pp. 2–7.

[34] H. White, J. Penney, A. Gibson, A. Szakay, and F. Cox, "Creak prevalence and prosodic context in Australian English," in *Annual Conference of the International Speech Communication Association (24th: 2023)*. International Speech Communication Association (ISCA), 2023, pp. 112–116.